

5

VIDEO CODING METHOD AND APPARATUS THEREOF

BACKGROUND OF THE INVENTION

Field of the Invention

10 [0001] The present invention is generally related to a technique for enhancing the quality of an image. More particularly, the present invention relates to a region-of-interest (ROI) video-coding algorithm based on fuzzy control method for a video encoder, for example, a H.263+ type video encoder.

15 Description of the Related Art

 [0002] The demand for applications of the digital video communication, such as videoconferencing and videophone, has increased considerably. However, the transmission rates over network are restricted, hence very low bit-rate video coding for such applications is an important technology to reduce the data rate of picture sequence
20 without losing much of its subjective quality. Most implementations of these standards give equal importance to each block. While different blocks within the same picture may be coded with different modes, no one block is more important than the other is. This model is not appropriate for any region-of-interest (ROI) application on video sequence. In H.263+ standard, the distortion weight parameter and the signal variance at
25 macro-block (MB) layer are adjusted to control the qualities at different regions. The blocks correspond to some focus areas are more important than the blocks in the background or unwanted areas. Allocating more bandwidth towards the quality of areas that user focuses on, while sacrificing background or unwanted areas quality is a better

coding strategy for video sequences like video conferencing. Except the ROI has more high quality, it may discard some background information to improve the encoding speed. Like maximum bit transfer (MBT), the background is always encoded with the coarsest quantization level as in. A region-based blurring algorithm to reduce bit-rate in very low bit-rate video coding is adopted. Another method improves quality at ROI significantly by three fixed factors to each ROI MBs and non-ROI MBs in order to enhance the quality of ROI regions, and reduce the bits for coding the background. The present invention can improve ROI quality adaptively according to fuzzy logic rate control and it is suitable for real time videoconferencing.

[0003] Fuzzy logic was first proposed by L.A. Zadeh working at Berkeley in 1965 and it is modeled after the natural way people arrive at solutions in three points. The first point: applying different solution methodologies to the same problem. The second point: applying more than one of our rules to the same problem at the same time. The third point: accepting a certain amount of imprecision, which is very important at helping us arrive at workable solutions. Obviously, normal rate control algorithms in different standard test models, such as TMN5, TMN8, and etc., are conformed to these three points. In each test models, there are particular mathematical solutions to determine the quantization parameters for each MB and a few inaccuracies are acceptable to estimate the bit rate for the next MB. It seems that a fuzzy logic control could play a suitable role in solving the rate control in video coding.

[0004] FIG. 1a shows a block diagram of a conventional feedback control system 100. This controller makes its decisions about what to do based on either a mathematical model of the process or a fixed set of mathematical relationship.

5 **[0005]** FIG. 1b shows a block diagram of a fuzzy logic control system 150. The fuzzy logic controller 150 uses as its guide a set of response rules established by the knowledgeable operators or system engineers. Referring to FIG. 1b, a quantizer 152 takes the data from a sensor 157 and converts the data into a format, which can be used by a fuzzy logic controller 153. The fuzzy logic controller 153 then performs
10 calculations to determine a fuzzy situation for that particular data.

[0006] To summarize, as the information highway has already begun, and with a limited transmission rate, a method for enhancing an image is needed. Currently, a region-of-interest (ROI) method that can improve an image's quality is already existed. However, the present solutions for the ROI methods still have barriers in the
15 performance. Therefore and for the foregoing reasons, there is a desperate need for a method or algorithm that is able to obtain a high quality video image.

SUMMARY OF THE INVENTION

[0007] The present invention is directed to a method and apparatus that satisfies
20 the need to enhance the quality of an image in applications such as videophone and videoconferencing. To achieve these and other advantages and in accordance with the purpose of the invention, as embodied and broadly described herein, a new method and apparatus based on region-of-interest (ROI) and fuzzy logic control are provided.

[0008] First, the method separates a plurality of region-of-interest regions from a
25 plurality of non-region-of-interest regions of an image. Then, an input from the region-of-interest regions is sent to a fuzzy logic controller, wherein the fuzzy logic controller is used for enhancing the quality of the region-of-interest regions and the overall quality of an output image.

5 **[0009]** In one preferred embodiment of the present invention, the input from the region-of-interest regions is calculated from a first control input and a second control input from the region-of-interest regions. Wherein, the first control input and the second control input comprise a first variance from a present (i)th macro-block and a variance difference, respectively. The variance difference is calculated by subtracting a second
10 variance of a previous (i-1)th macro-block from the first variance and then dividing by the first variance. The (i)th macro-block and the (i-1)th macro-block represent a sequence of macro-block within one of the region-of-interest regions and the (i-1)th macro-block is a previous macro-block of the (i)th macro-block.

[0010] In another preferred embodiment of the present invention, the fuzzy logic
15 control includes a methodology to convert the control inputs to fuzzy predicates

[0011] In another preferred embodiment of the present invention, the fuzzy logic control includes a controlling function to calculate a linguistic membership function for determining a fuzzy situation of the main control input. The controlling function uses center of area (COA) method to determine the linguistic membership function.

20 **[0012]** In another embodiment of the present invention, the fuzzy logic control includes a plurality of lookup tables for making a decisional level and producing a weighted factor to emphasize the qualities of one of the region-of-interest regions.

[0013] In yet another embodiment of the present invention, the lookup tables comprise a plurality of scaled lookup tables for providing a priority-like quality for one
25 of the region-of-interest regions. Wherein, the scaled lookup tables are formed by using a one-fixed and one-various membership function.

[0014] To summarize, a fuzzy controlled ROI video coding is provided. The fuzzy controlled ROI video coding has the capability of adjusting the output quality of

5 an image adaptively. The approach can enhance the quality of ROI easily, maintain the constant bit-rate to avoid buffer overflow, and achieve good quality easily with fewer bit-rates than previous works. The multiple ROI video coding can also enhance each ROI's output quality significantly without complex computation.

[0015] It is to be understood that both the foregoing general description and the
10 following detailed description are exemplary, and are intended to provide further explanation of the invention as claimed.

BRIEF DESCRIPTION OF THE DRAWINGS

[0016] The accompanying drawings are included to provide a further
15 understanding of the invention, and are incorporated in and constitute a part of this specification. The drawings illustrate embodiments of the invention and, together with the description, serve to explain the principles of the invention.

[0017] FIG. 1a illustrates a conventional feedback control algorithm.

[0018] FIG. 1b illustrates a conventional fuzzy logic control algorithm.

20 [0019] FIG. 2 illustrates one embodiment of the present invention showing a block diagram of region-of-interest video coding by fuzzy logic control algorithm.

[0020] FIG. 3 illustrates one version of a variance i subsets of the fuzzy logic control device as shown in FIG. 2.

[0021] FIG. 4 illustrates one version of a variance change Δi subsets of the fuzzy
25 logic control device as shown in FIG. 2.

[0022] FIG. 5 illustrates one version of a fuzzy output lookup table of the fuzzy logic control device as shown in FIG. 2.

5 **[0023]** FIG. 6 illustrates one version of a one-fixed and one-various membership function.

[0024] FIG. 7 illustrates one comparison of different methods for Carphone sequence at 64 kbits/sec for 100 frames.

[0025] FIG. 8 illustrates one comparison of different methods for Claire
10 sequence at 32 kbits/sec for 150 frames.

[0026] FIG. 9 illustrates one comparison of different methods for Foreman sequence at 64 kbits/sec for 150 frames.

[0027] FIG. 10 illustrates one comparison of multiple region-of-interest for News sequence at 64 kbits/sec for 150 frames.

15

DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0028] The present invention now will be described more fully hereinafter with reference to the accompanying drawings, in which preferred embodiments of the invention are shown. This invention may, however, be embodied in many different
20 forms and should not be construed as limited to the embodiments set forth herein; rather, these embodiments are provided so that this disclosure will be thorough and complete, and will fully convey the scope of the invention to those skilled in the art. Like numbers refer to like elements throughout.

[0029] To begin with, a region-of-interest video coding by fuzzy control,
25 consisted of two main components: (1) a region-of-interest, and (2) a fuzzy control. Referring to FIG. 2, a region-of-interest includes segmentation 302. Whereas a fuzzy logic controller 320 includes: a differential variance calculator 303; a quantizer 304; fuzzy subsets 305; a fuzzy controller 306; a fuzzy variance operator 307; a weighted

5 defuzzifier 308; and a fuzzy lookup table 309. In addition, a H.263+ video encoder and a virtual buffer are also included for an overall coding system.

[0030] Also referring to FIG. 2, a fuzzy logic controller 320 enhances the quality of region-of-interest according to a variance σ_i 332 and a variance difference $\Delta\sigma_i$ 334. After a frame 301 is input, the segmentation 302, such as face detection and motion
10 detection, are used to separate the frame 301 into region-of-interest (ROI) regions 330 and non-ROI regions 331. The macro-blocks in non-ROI region 331 are sent directly to a QP selection 310 in rate control without adjusting any parameters. The variance difference $\Delta\sigma_i$ 334 in the i -th macro-block of one of the ROI regions 330 is calculated from σ_i 332 and σ_i' 333, where σ_i 332 and σ_i' 333 are variances of the current and the
15 previous i -th MB, respectively. The variance difference $\Delta\sigma_i$ 334 and the current MB variance σ_i 332 are the two inputs to apply the fuzzy logic method and ω_{σ_i} 335 is a fuzzy output to be the weighted factor of input.

[0031] FIG. 3 and FIG. 4 are the graphical representations of σ_i 332 and $\Delta\sigma_i$ 334, respectively. Referring to FIG. 3 and FIG. 4, the notations, which are qualitative
20 statements of linguistic sets, LN 351 and 401, SN 352 and 402, ZE 353 and 403, LP 354 and 404, and SP 355 and 405 are “Large Positive”, “Small Positive”, “Zero”, “Small Negative” and “Large Negative”, respectively. The notations of FIG. 3 are the same as that of FIG. 4 except all the σ_i 332 are positive and the most variances σ_i 334 of each MB center on ZE 303 in the statistics. FIG. 4 shows the subsets of the variance
25 difference $\Delta\sigma_i$ 334, which is defined as $\Delta\sigma_i = (\sigma_i - \sigma_i') / \sigma_i$

[0032] Referring to FIG. 4, most $\Delta\sigma_i$ 334 are concentrated in [-10, +10] in the statistics. Next, the quantizer 304 takes the σ_i 332 and $\Delta\sigma_i$ 334 into the fuzzy subsets 305 and convert their degrees into fuzzy predicates such as LN 351, SN 352, ZE 353, LP

5 354, and SP 355. The fuzzy controller 306 then calculates the linguistic membership function by the quantized σ_i 332 and $\Delta\sigma_i$ 334, and utilizes the center of area (COA) method to determine the fuzzy situation. After the calculations, each $\sigma_i/\Delta\sigma_i$ pair has a corresponding main control input value. The decision table is stored in memory in the form of a fuzzy lookup table 309 as shown in FIG. 5. The weighted defuzzifier 308
 10 takes the two situations of $\sigma_i/\Delta\sigma_i$ into account according to the fuzzy lookup table 309 and ω_{oi} 335, the weighted factor, is outputted to emphasize the ROI 330 macro-blocks' qualities.

[0033] In one embodiment of the present invention, a set of different output fuzzy tables is scaled by the original output fuzzy in order to have different priorities to
 15 different ROI regions 330. FIG. 6 describes a one-fixed and one-various membership function, which is used to utilize and distinguish the different ROI 330 from each ROI priority. The weighted factors are calculated by the fuzzy rule and given to each MB in the H.263+ video encoder 311.

[0034] As an experimentation for one embodiment of the present invention
 20 shows the embodiment of the present invention has a better performance than other existing methodologies. In the experimental results, three sequences: Carphone; Claire; and Foreman are tested. In order to define the ROI regions in a frame, a face detection is used to select ROI automatically. Four different methods in the test sequences are compared. The four different methods are: coding a frame without ROI (WR), coding
 25 the ROI regions by multiplying a weighted factor (WA) α , coding the ROI regions by three factors (TF), and the present invention (Fuzzy). The four different methods are all set to the similar average bit-rate. In an implementation, QP is set to 5 and 3 for I-frame and P-frame at target bit-rate 64 kbits/sec, and 15 and 13 for I-frame and P-frame at

5 target bit-rate 32 kbits/sec, respectively. In WA, the weighted factor is set to be 450. In TF, the three factors are set to be 450, 2, and 10, respectively. In order to compare the other two methods in similar weights, ZE_{13} is set to be 450 and $LP_1 \sim LN_{25}$ are set to be in 350~550.

10 [0035] As illustrated from FIG. 7 to FIG. 10, the embodiment of the present invention has a better PSNR of ROI in the similar bit-rates compared to the other methods. Since both of WA and TF enhance the ROI quality by fixed factors, the two methods cannot adjust the weighted factor when the complexity of each MB changes rapidly. To summarize, the embodiment of the present invention obtains better quality in ROI regions and less skipping frames even with lower bit-rate.

15 [0036] The present invention is suitable in any image processing. It is particular useful for real-time video coding. Accordingly, the present invention can enhance the quality of ROI easily and maintain the constant bit-rate to avoid buffer overflow. It can achieve good quality easily with fewer bit-rates than previous works. The multiple ROI video coding can also enhance each ROI's quality significantly without complexity
20 computation.

[0037] It will be apparent to those skilled in the art that various modifications and variations can be made to the structure of the present invention without departing from the scope or spirit of the invention. In view of the foregoing, it is intended that the present invention cover modifications and variations of this invention provided they
25 fall within the scope of the following claims and their equivalents.